1. (6 points) Classify the following variables as either categorical or numerical by checking the correct box, if it is a numerical variable, further classify the variable as either discrete or continuous:

	Variable	Categorical	Numerical	Discrete	Continuous
A	Hair Color	✓			
B	Frozen Food Brand	✓			
C	Number of students in a classroom		✓	✓	
D	Your age		✓		✓

2.(8 points) To study the effects of exercise on the grades of college students, a researcher wishes to compare the grade point averages of students at randomly selected colleges across the United States. The researcher selects students at random and after interviewing them to find out who exercises and who does not, chose 644 students of each (exercisers and non-exercisers). The researcher made sure the two groups of 644 were similar in racial composition, gender, major, and every subject had accumulated at least 120 units towards graduation. There were a total of 1,288 students in the study from approximately 40 colleges, their overall GPA was 3.22. The average GPA for the students who exercised was 3.34 and the standard deviation was .36.

a. What is the "treatment"?
Exercising (not assigned by researcher)

b. What is the response or outcome variable?
GPA

c. Is this an observational study or an experiment?
observational study

d. From this study, an example of a sample statistic is:
3.34 - Average G.P.A for the 644 students who exercised
SD = .36

e. What is the population of interest?
all college students

f. What is the parameter of interest in this study?
average GPA of all exercising college students (as compared with non-exercising college students)

3. (6 points) The next questions refer to the list {-4, -9, 0, -3}.

a. What is the mean of this list?

$$\text{mean} = \frac{-4-9+0-3}{4} = \boxed{-4}$$

b. What is the standard deviation of this list?

$$SD = \sqrt{\frac{(-4+4)^2+(-9+4)^2+(0+4)^2+(-3+4)^2}{4}}$$

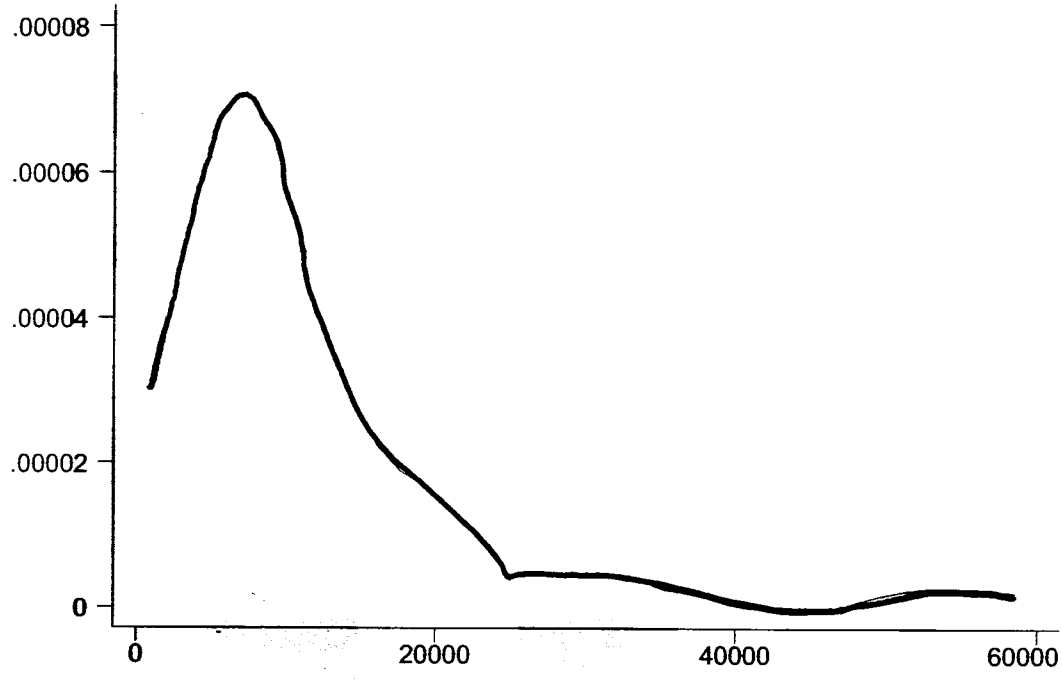
$$SD = \sqrt{\frac{0+25+16+1}{4}}$$

$$SD = \sqrt{\frac{42}{4}}$$

$$SD = \sqrt{10.5}$$

$$\boxed{SD \approx 3.2404}$$

4. (5 points) Here is a histogram, please assume it was correctly drawn

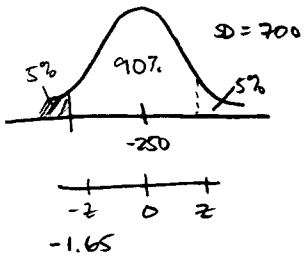


Please answer the following questions about this histogram:

	True	False	Statement
A	X		The mean is greater than the median
B	X		This histogram is right skewed
C		X	This histogram is left skewed
D	X		The total area is 1.0 or 100%
E	X		If the X axis represents dollars, the implied Y axis label is percent per dollar

The last three questions use information from this statement, but each question is separate (i.e. you can get the first one wrong and it won't affect the others): A recent study showed that the gambling income of adults age 21 and over in the United States from all forms of legalized gambling (e.g. lottery, video poker, horse racing, casinos) is normally distributed with a mean of -250 dollars (a loss) and a standard deviation of $700. SHOW YOUR WORK FOR FULL CREDIT.

5. It is believed that the gamblers with the largest losses, that is those with the lowest 5% of gambling income, should be considered gambling "addicts" and given some kind of treatment. How much money does a gambling adult need to lose to be considered an "addict"? (5 points)



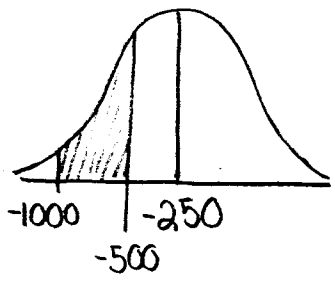
where area = 90%, $z = 1.65$

$$z = \frac{x-\mu}{\sigma} \qquad -1.65 = \frac{x-(-250)}{700}$$

$$-1155 = x + 250$$

$$x = -1405$$

6. What percentage of adults age 21 and over had gambling losses of at least $500 but not more than $1000? (7 points) $\bar{X}=-250$ SD=700



$$\frac{-500+250}{700}=-0.36 \rightarrow 27.37\%$$

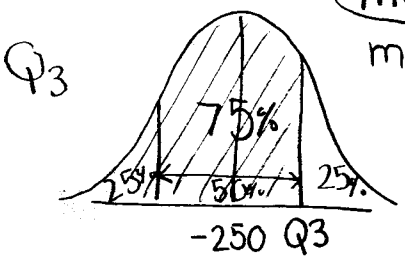
$$\frac{-1000+250}{700}=-1.07 \rightarrow 70.63\%$$

$$\frac{70.63-27.37}{2}=21.63\%$$

7. What is the median gambling income? What is the interquartile range for gambling income? (8 points)

median=-250

median=$\bar{X}$ in a normal curve

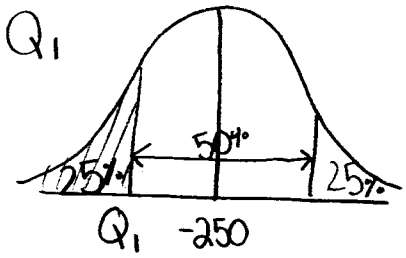


50% → 0.65

$$\frac{X+250}{700}=0.65$$

IQR=Q_3-Q_1

$Q_3 \rightarrow X=205$



50% → -0.65

$$\frac{X+250}{700}=-0.65$$

$Q_1 \rightarrow X=-705$

IQR = $Q_3 - Q_1$
= 205+705
= 910